Amendments to the Claims

D1

1. (Currently amended) A method of predicting a biological activity of a test compound, comprising:

obtaining spectral data for the test compound and for a training set of compounds having known biological activities;

segmenting the spectral data of the training set of compounds into bins;

scaling the segmented spectral data of the training set of compounds prior to establishing a spectral data-activity relationship [deriving a pattern of spectral data associated with the biological activity];

weighting the segmented spectral data of the training set of compounds prior to establishing the spectral data-activity relationship;

establishing the spectral data-activity relationship between the known biological activities and the segmented, scaled and weighted spectral data of the training set of compounds using computer implemented pattern recognition; [deriving the pattern of spectral data associated with the biological activity from the spectral data of the training set of compounds; and]

segmenting the spectral data of the test compound into bins; and

predicting the biological activity of the test compound from its segmented spectral data using the spectral data-activity relationship. [by detecting similarities between the pattern of spectral data associated with the biological activity and a pattern of spectral data for the test compound.]

2. (Previously amended) The method of claim 1, wherein the spectral data are obtained without first correlating the spectral data with corresponding structural features.

3. (Currently amended) The method of claim 1, wherein the [pattern of spectral data associated with a biological activity is derived] spectral data-activity relationship is established without first correlating the spectral data with corresponding structural features.

4. (Cancelled)

5. (Currently amended) The method of claim 1 [4], wherein the spectral data of the test compound is segmented into substantially the same bins as the spectral data of the training set. [the pattern of spectral data of the test compound is obtained by segmenting the spectral data of the test compound into substantially the same sub-spectral units into which the spectral data of the training set is segmented.]

6. (Original) The method of claim 1, wherein the spectral data is one type of spectral data.

7. (Previously amended) The method of claim 6, wherein the spectral data is one of nuclear magnetic resonance, mass spectral, infrared, ultraviolet-visible, fluorescence, or phosphorescence data.

8. (Original) The method of claim 1, wherein the spectral data is a composite of different types of spectral data.

9. (Previously amended) The method of claim 8, wherein the composite comprises two or more of the group consisting of nuclear magnetic spectroscopy (NMR), mass spectroscopy (MS), infrared (IR) spectroscopy, and ultraviolet-visible (UV-Vis) spectroscopy.

10. (Previously amended) The method of claim 1, wherein the spectral data of the test compound is segmented into substantially the same spectral sub-units as the spectral data of the training set of compounds to produce the spectral pattern for the test compound.

11. (Previously amended) The method of claim 1, wherein scaling comprises auto-scaling.

12. (Cancelled)

13. (Currently amended) The method of claim 1 [12], wherein weighting comprises Fisher-weighting.

14. (Previously amended) The method of claim 1, wherein establishing and predicting [detecting similarities between the pattern of spectral data associated with a biological activity of the training set and the pattern of spectral data for the test compound] comprises statistical pattern recognition.

15. (Cancelled)

16. (Cancelled)

17. (Currently amended) The method of claim 1, wherein establishing and predicting [the method] comprise[s] artificial intelligence pattern recognition.

18. (Currently amended) A computer implemented method for predicting a biological activity of a test compound, comprising:

receiving spectral data for a test compound as input;

receiving spectral data and endpoint data of a training set of compounds having known biological activities as input;

segmenting the spectral data of the training set of compounds into sub-spectral units;

scaling the segmented spectral data of the training set of compounds;

weighting the segmented spectral data of the training set of compounds;

establishing a spectral data-activity relationship between the segmented, scaled and weighted spectral data and the known biological activities of the training set of compounds using pattern recognition [detecting a pattern of spectral data associated with the biological activity]; and

predicting the biological activity of the test compound using the spectral data-activity relationship. [by comparing the pattern of spectral data associated with the biological activity to the spectral data of the test compound to determine whether the spectral data of the test compound is similar to the spectral pattern associated with the biological activity of the training set and the test compound is predicted to share the biological activity.]

19. (Currently amended) The computer implemented method of claim 18, wherein [comparing comprises comparing] establishing and predicting are performed with a statistical pattern recognition program.

20. (Currently amended) The computer implemented method of claim 19, wherein the spectral data for the test compound is segmented into substantially identical sub-spectral units as the training set spectral data. [, so that a signal within an individual sub-spectral unit is compared to the corresponding sub-spectral unit of the pattern.]

21. (Previously amended) The computer implemented method of claim 18, wherein the spectral data are selected from the group consisting of nuclear magnetic resonance data, mass spectral data, infrared data, ultraviolet-visible data, fluorescence data, phosphorescence data, and composites of two or more such spectral data.

22. (Currently amended) The computer implemented method of claim 21, wherein the spectral data-activity relationship comprises canonical variate factors. [pattern associated with the biological activity of the training set is a set of canonical variate factors, and the spectral data for the test compound are compared to the canonical variate factors of the training set spectral data.]

23. (Currently amended) The computer implemented method of claim 22, wherein the biological activity is binding affinity to a hormone receptor. [, and the canonical variate factors include peaks in sub-spectral units that are associated with hormone receptor binding of a pre-selected affinity.]

24. (Previously amended) The computer implemented method of claim 23, wherein the spectral data comprise nuclear magnetic resonance data and mass spectral data.

25. (Original) A computer readable medium having stored thereon instructions for performing the actions of claim 1.

26. (Original) A computer readable medium having stored thereon instructions for performing the actions of claim 18.

Claims 27-54 (Cancelled)

55. (Previously amended) The method of claim 1, wherein the spectral data comprises calculated spectral data.

56. (Previously amended) The computer implemented method of claim 18, wherein the spectral data comprises calculated spectral data.

57. (Previously amended) The computer implemented method of claim 56, wherein the calculated spectral data comprises calculated nuclear magnetic resonance data.

58. (Previously amended) The computer implemented method of claim 57, wherein the calculated nuclear magnetic resonance data comprises calculated ^{13}C NMR data.

Claims 59-64 (Cancelled)

65. (Currently amended) The method of claim 1 further comprising:

predicting a second biological activity of the test compound using a second spectral data-activity relationship, the second spectral data-activity relationship established between the scaled and weighted spectral data of a second training set of compounds and the known biological activities of the second training set compounds. [by comparing the spectral data of the test compound to a second pattern of spectral data associated with the second biological activity to determine if the test compound shares the second biological activity, the second pattern derived using scaled spectral data and known endpoints of a training set of compounds for the second biological activity.]

66. (Previously added) The method of claim 1, wherein the spectral data comprises ^{1}H, ^{13}C, ^{15}N, ^{17}O, ^{19}F, ^{31}P or ^{35}S NMR data.

67. (Previously added) The method of claim 66, wherein the spectral data comprises calculated spectral data.

68. (Previously added) The method of claim 57, wherein the calculated nuclear magnetic resonance data comprises ^{1}H, ^{13}C, ^{15}N, ^{17}O, ^{19}F, ^{31}P or ^{35}S NMR data.

69. (Currently amended) A method for predicting a biological property of a test compound, comprising:

[providing biological activity data for a plurality of compounds;]

providing spectral data for [the plurality] a training set of compounds having known biological activities;

providing spectral data for the test compound;

segmenting the spectral data for the test compound [and the plurality of compounds] into bins;

segmenting the spectral data for the training set of compounds into bins;

scaling the spectral data of the training set [plurality] of compounds in the bins;

weighting the spectral data of the training set [plurality] of compounds in the bins;

establishing a spectral data-activity relationship between the known biological activities and the segmented spectral data of the training set of compounds using computer implemented statistical pattern recognition[detecting a pattern of spectral data of the plurality of compounds in the bins that is correlated with the biological activity]; and

using the spectral data-activity relationship to predict the biological activity of the test compound from its segmented spectral data [detecting similarities between the pattern correlated with the biological activity and the spectral data of the test compound to determine if the molecule shares the biological activity].

70. (Previously added) The method of claim 69, wherein scaling comprises autoscaling.

71. (Previously added) The method of claim 70, wherein weighting comprises Fisher-weighting.

72. (Previously added) The method of claim 69, wherein weighting comprises Fisher-weighting.

73. (Previously added) The method of claim 69, wherein scaling comprises variance scaling.

74. (Currently amended) The method of claim 69, wherein the spectral data-activity relationship [pattern detected in the bins] comprises a set of canonical variate factors.

75. (Previously added) The method of claim 69, wherein the spectral data comprises nuclear magnetic resonance, mass spectral, infrared, ultraviolet-visible, fluorescence, or phosphorescence data.

76. (Previously added) The method of claim 75, wherein the spectral data comprises ^{13}C NMR data.

77. (Previously added) The method of claim 76, wherein the ^{13}C NMR data comprises calculated ^{13}C NMR data.

78. (Previously added) The method of claim 75, wherein the spectral data comprises a composite of two or more types of spectral data.

79. (Previously added) The method of claim 78, wherein the composite comprises ^{13}C NMR data and EI-MS data.

80. (Previously added) The method of claim 69, wherein detecting a pattern comprises statistical pattern recognition.

81. (Previously added) The method of claim 76, wherein segmenting into bins comprises segmenting the ^{13}C NMR data into bins having a width from 0.5 ppm to 5.0 ppm.

82. (Currently amended) A computer implemented method for predicting the biological activity of a test compound, comprising:

receiving as input spectral data for a test compound;

receiving as input training set data, the training data comprising the spectral data and biological activities of a training set of compounds;

segmenting the spectral data of the training set into bins;

autoscaling the spectral data of the training set;

Fisher-weighting the spectral data of the training set;

using a pattern recognition method to establish a spectral data-activity relationship that classifies compounds of the training set into two or more endpoint classes from the [detecting a pattern of] segmented, autoscaled and Fisher-weighted spectral data of the training set and the known biological activities [characteristic of an endpoint class of compounds in the training set of compounds];

segmenting the spectral data for the test compound into bins; and

using the spectral data-activity relationship to predict the test compound's endpoint class. [predicting the biological activity of the test compound by detecting similarities between the segmented spectral data of the test compound and the pattern of spectral data characteristic of the endpoint class to predict whether the test compound's is also in the endpoint class.]

83. (Currently amended) The computer implemented method of claim 82, wherein [detecting] the spectral data-activity relationship comprises canonical variate factors for the bins. [a pattern of spectral data associated with an endpoint class of compounds in the training set of compounds comprises calculating canonical variate factors for the bins.]

84. (Previously added) The computer implemented method of claim 82, wherein the spectral data of the test compound and the spectral data of the training set of compounds comprise spectral data selected from the group consisting of nuclear magnetic resonance data, mass spectral data, infrared data, ultraviolet-visible data, fluorescence data, phosphorescence data, and composites thereof.

85. (Previously added) The computer implemented method of claim 84, wherein the spectral data of the test compound and the spectral data of the training set of compounds comprises a composite of two or more types of spectral data and the spectral data is normalized to yield structure descriptors of similar magnitude from each type of spectral data.

86. (Previously added) The computer implemented method of claim 84, wherein the spectral data comprises ^{1}H, ^{13}C, ^{15}N, ^{17}O, ^{19}F, ^{31}P or ^{35}S NMR data.

87. (Previously added) The method of claim 86, wherein the spectral data is calculated.

88. (Previously added) The computer implemented method of claim 82, wherein the spectral data of the test compound and the spectral data of the training set of compounds comprise ^{13}C NMR data and segmenting into bins comprises dividing the ^{13}C NMR data into sub-spectral units having a width from 0.5 ppm to 5.0 ppm.

89. (Previously added) The computer implemented method of claim 88, wherein the ^{13}C NMR data comprises calculated ^{13}C NMR data.

90. (Previously added) A computer readable medium having stored thereon instructions for performing the actions of claim 82.

91. (New) The method of claim 82 further comprising validating the spectral data-activity relationship.

92. (New) The method of claim 91, wherein validating comprises leave-one-out cross-validation.

93. (New) The method of claim 82, wherein the pattern recognition method is a statistical pattern recognition method.